ELSEVIER

Journal of Computational and Applied Mathematics 53 (1994) 33–42

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Strong convergence of almost simultaneous block-iterative projection methods in Hilbert spaces

Dan Butnariu *, Yair Censor [1]

Department of Mathematics and Computer Science, University of Haifa, 31905 Haifa, Israel

Received 19 May 1992; revised 4 November 1992

Abstract

We prove strong convergence of a class of block-iterative projection methods for finding a common point of a finite family of closed convex subsets in a Hilbert space.

Keywords: Convex feasibility problem; Block-iterative projection method; Strong convergence; Weak convergence; Uniformly convex set; Boundedly compact set

1. Introduction

Let $\{Q_i \mid 1 \leqslant i \leqslant m\}$ be a finite family of closed convex sets in a Hilbert space $\mathscr{H}$ and assume that $Q := \bigcap_{i=1}^{m} Q_i$ is nonempty. The *convex feasibility problem* (CFP, for short) is to find an element $x^* \in Q$. The *block-iterative projection* (BIP) algorithmic scheme for solving the CFP, proposed in [1] in the Euclidean space $\mathbb{R}^n$ setting and further studied in [4,7,9], iteratively generates a sequence as follows. Choose an *initial point* $x^0 \in \mathscr{H}$ and, for each $k \in \mathbb{N}$, do

$$x^{k+1} = x^k + \lambda_k \sum_{i \in I} w_k(i)\big(P_i(x^k) - x^k\big), \tag{1}$$

where $I := \{1, 2, \ldots, m\}$, $P_i(x^k)$ is the orthogonal projection of x^k onto the set Q_i, $w_k : I \to \mathbb{R}_+$ is a *weight function* (i.e., $\sum_{i \in I} w_k(i) = 1$) and $\lambda_k \in \mathbb{R}_+$ are *relaxation parameters*. In what follows we assume that there exist two real numbers ϵ_1 and ϵ_2 such that, for all $k \in \mathbb{N}$,

$$0 < \epsilon_1 \leqslant \lambda_k \leqslant \epsilon_2 < 2. \tag{2}$$

* Corresponding author. e-mail: rsma606@haifauvm.bitnet.

[1] The work of this author was supported by NIH Grant HL-28438, while visiting the Medical Image Processing Group (MIPG) at the Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, PA, United States.

SSDI 0377-0427(92)00123-O

Our objective in this paper is to study the strong convergence of BIP methods in Hilbert spaces. It is known (cf. [3,4,8]) that, under quite mild conditions on the weight functions and the relaxation parameters, BIP methods generated sequences *converge weakly* to points in Q, regardless of the choice of the initial point. Ensuring *strong convergence* of such sequences to points in Q is somewhat more difficult and usually demands additional conditions on the sets Q_i themselves.

Classical strong convergence theorems for BIP methods are due to von Neumann [18] and Halperin [11] for sets Q_i which are closed linear subspaces of $\mathscr{H}$ and Gubin et al. [10] for sets Q_i which are closed halfspaces in $\mathscr{H}$ or are uniformly convex or satisfy $\mathrm{Int}(Q) \neq \emptyset$. These results guarantee strong convergence of particular sequential BIP methods. Recall that a BIP method is *sequential* if all weight functions w_k are Kronecker vectors in $\mathbb{R}^m$.

It was repeatedly observed (see [1,4,8,12,15]) that convergence of sequential BIP methods is slow. This led to the question whether, and under what conditions, nonsequential BIP methods converge strongly and if their convergence is faster. The first strong convergence result for nonsequential BIP methods, due to Pierra [15], shows that BIP methods with *uniformly distributed* weight functions (i.e., with $w_k(i) = 1/m$, $i \in I$) and appropriately chosen relaxation parameters strongly converge in Hilbert spaces provided that the CFP satisfies conditions similar to those in [10]. De Pierro and Iusem [8] extended Pierra's result by showing that simultaneous BIP methods with constant sequences of relaxation parameters strongly converge to points in Q when one of the sets Q_i is compact. Recall [12] that a BIP method is called *simultaneous* if, for all $k \in \mathbb{N}$, $w_k(i) = w(i) > 0$, $i \in I$, where $w(\cdot)$ is some fixed weight function. It follows from [15] that simultaneous BIP methods are equivalent to cyclically controlled sequential BIP methods in a product Hilbert space. Thus, one should not expect that simultaneous BIP methods would have better rates of convergence [13] or initial speeds [5] than their sequential counterparts.

Computational experience with BIP methods has shown that conveniently varying the weight functions along the iterative process of generating BIP sequences may eventually improve the initial speed of convergence, see, e.g., [1,4,9,12]. Also, BIP methods with conveniently chosen weight functions lend themselves to parallel implementations which produce practical computational gains on properly chosen computing architectures. These facts lead to the question whether nonsequential and nonsimultaneous BIP method generated sequences still converge strongly to solutions of the CFPs from which they are derived.

A comprehensive analysis of strong convergence of not necessarily simultaneous BIP methods *in general Hilbert spaces* is due to Ottavy [14]. He shows that, if the sets Q_i satisfy some topological conditions, BIP generated sequences whose weight functions and relaxation parameters are chosen at each iterative step according to specific rules (see (7) and (8) below) strongly converge to points in Q. Ottavy's rules affect the maximal size of the allowed relaxation parameters and, therefore, may limit the initial speeds of the BIP generated sequences, see Section 3. Also, these rules implicitly require computation of a large number of projections $P_i(x^k)$, a task which is computationally costly when the sets Q_i have complex geometries.

In this note we prove strong convergence of "almost simultaneous" BIP methods. A BIP method is called *almost simultaneous* [4] if the sequence $\{w_k \mid k \in \mathbb{N}\}$ has a convergent subsequence whose limit in $\mathbb{R}^I$ is a positive weight function w_*. The class of almost simultaneous BIP schemes does not include sequential methods, so our analysis does not cover but rather complements earlier work on strong convergence of sequential BIP methods in [10,11,18]. On the other hand, our analysis does cover a large class of BIP methods not dealt with before such as *pseudo-periodical* BIP procedures. These are BIP schemes such that, for some fixed integer $q \geqslant 2$ and for some fixed positive weight function w, $w_k = w$ whenever k is divisible by q while the other weight functions are chosen arbitrarily. Thus, our main result, Theorem 1 of Section 2, improves upon the results in [8,15] by guaranteeing strong convergence of BIP method generated sequences under less restrictive controls and under equivalent or weakened conditions on the sets Q_i. Also, Theorem 1 complements the convergence results in [14] and extends [4, Theorem 4.4].

BIP methods for solving CFPs have far-reaching applications in the field of image recovery [16], the theory of learning in neural nets [19] and computerized tomography [6,7]. Also, strongly convergent BIP methods can be instrumental in solving convex optimization [17] and optimal control problems [10,15]. In many such applications strongly convergent BIP method generated sequences can be constructed with weight functions w_k which may change dynamically based on current and past iterates, see [6]. The freedom of choosing the weight functions as guaranteed by Theorem 1 provides a tool of dynamic parallel processing and, thus, it could contribute to speeding up computations as shown in [4]. The fact that various choices of weight functions and relaxation parameters may lead to BIP generated sequences with substantially different initial speeds and rates of convergence can be observed in practice. A fully reliable rule of dependence of the rates of convergence of BIP schemes upon the choice of the weight functions or of the relaxation parameters is still missing. However, the wide range of such choices allowed by Theorem 1 permits construction of BIP procedures with increased initial speed of convergence (see Section 3).

2. A strong convergence theorem

In this section we consider the CFP in a Hilbert space $\mathscr{H}$ and assume that $\{x^k \mid k \in \mathbb{N}\}$ is a sequence generated by a BIP method starting from an arbitrary initial point $x^0 \in \mathscr{H}$. According to [10] a set Q_i is *uniformly convex* if there exists a function $\delta :]0, \infty[\to]0, \infty[$ such that whenever $x, y \in Q_i$, $x \neq y$, $z \in \mathscr{H}$ and $\| z - \frac{1}{2}(x+y) \| < \delta(\| x - y \|)$, we have $z \in Q_i$. A set Q_i is called *boundedly compact* if its intersection with any closed ball in $\mathscr{H}$ is compact. Note that a closed set Q_i is boundedly compact iff any bounded sequence included in Q_i has a strongly convergent subsequence.

Theorem 1. *If a BIP method is almost simultaneous with relaxation parameters as in* (2) *and if any of the following conditions is satisfied*:

(A) *there exists $i_0 \in I$ such that $Q_{i_0} \cap \operatorname{Int}[\bigcap_{i \neq i_0} Q_i] \neq \emptyset$*;

(B) *all, except for possibly one, of the sets Q_i are uniformly convex*;

(C) *each Q_i is a halfspace (i.e., $Q_i = \{x \in \mathscr{H} \mid \langle x, c_i \rangle \leqslant \beta_i\}$ for some $c_i \in \mathscr{H}$ and for some $\beta_i \in \mathbb{R}$)*;

(D) *at least one set Q_i is boundedly compact*;

(E) *$\mathscr{H}$ is finite-dimensional*;

then any sequence $\{x^k \mid k \in \mathbb{N}\}$ generated by this method is strongly convergent in $\mathscr{H}$ to a point in $Q = \bigcap_{i=1}^m Q_i$.

The proof of this result consists of a sequence of lemmas which are given below. In [12] the *simultaneous projection operator* $\mathrm{P}_w : \mathscr{H} \to \mathscr{H}$, corresponding to the weight function $w : I \to \mathbb{R}_+$, was defined by

$$\mathrm{P}_w(x) = \sum_{i \in I} w(i) P_i(x),$$

and its following properties were established.

Lemma 2. (i) *If* $x,\ y \in \mathscr{H}$, *then* $\|\mathrm{P}_w(x) - \mathrm{P}_w(y)\| \leqslant \|x - y\|$.
(ii) *If* $x \in \mathscr{H}$ *and* $z \in Q$, *then* $\langle \mathrm{P}_w(x) - x,\ \mathrm{P}_w(x) - z\rangle \leqslant 0$.
(iii) *If* $x,\ y \in \mathscr{H}$ *then* $\langle \mathrm{P}_w(x) - \mathrm{P}_w(y),\ x - y\rangle \geqslant 0$.

Let $\{x^k \mid k \in \mathbb{N}\}$ be a sequence in $\mathscr{H}$ generated by an almost simultaneous BIP method with relaxation parameters satisfying (2). Then, applying [4, Corollary 4.3], we obtain the following lemma.

Lemma 3. *If* $z \in Q$ *and* $k \in \mathbb{N}$, *then* $\|x^{k+1} - z\| \leqslant \|x^k - z\|$.

Lemma 3 implies that the nonnegative sequence $\{\rho_k \mid k \in \mathbb{N}\}$ with $\rho_k := d_Q(x^k)$, where $d_Q(x^k)$ denotes the distance of x^k to the set Q, is nonincreasing and, hence, convergent. We denote $\rho := \lim_{k\to\infty}\rho_k$. Also by Lemma 3, since $Q \neq \emptyset$, there exists $z \in Q$ such that the sequence $\{x^k \mid k \in \mathbb{N}\}$ is contained in the closed ball centered at z whose radius is $\|x^0 - z\|$. Thus, the sequence $\{x^k \mid k \in \mathbb{N}\}$ is bounded in $\mathscr{H}$.

Let $\{w_{k_p} \mid p \in \mathbb{N}\}$ be a subsequence of $\{w_k \mid k \in \mathbb{N}\}$ such that $w_* = \lim_{p\to\infty} w_{k_p}$ exists in $\mathbb{R}^I$ and is positive. Such a subsequence exists since the BIP method is almost simultaneous. The corresponding subsequence $\{x^{k_p} \mid p \in \mathbb{N}\}$ is bounded and, therefore, it includes a weakly convergent subsequence which we again denote by $\{x^{k_p} \mid p \in \mathbb{N}\}$. Let x^* be the weak limit of $\{x^{k_p} \mid p \in \mathbb{N}\}$. Our aim is to show that $x^* \in Q$ and that $\{x^k \mid k \in \mathbb{N}\}$ converges strongly to x^*. To this end, we denote $a_k := \|\mathrm{P}_{w_k}(x^k) - x^k\|$, $k \in \mathbb{N}$, and show the following.

Lemma 4. *The series* $\sum_{k=1}^{\infty} a_k^2$ *converges.*

Proof. From [4, Proposition 3.3] we get that, for each $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} \lambda_i(2 - \lambda_i) a_i^2 \leqslant \rho_0^2 - \rho_{n+1}^2.$$

According to (2), $\lambda_i(2 - \lambda_i) \geqslant \epsilon_1(2 - \epsilon_2)$ when $1 \leqslant i \leqslant n$, and, thus,

$$\sum_{i=1}^{n} a_i^2 \leqslant (\rho_0^2 - \rho_{n+1}^2)[\epsilon_1(2 - \epsilon_2)]^{-1}.$$

Letting $n \to \infty$ and taking into account that the sequence $\{\rho_n \mid n \in \mathbb{N}\}$ converges to ρ, the convergence of $\sum_{k=1}^{\infty} a_k^2$ results. □

The next result is a consequence of Lemma 4.

Lemma 5. *The functional* $\phi_* : \mathscr{H} \to \mathbb{R}_+$, *defined by*

$$\phi_*(x) := \sum_{i \in I} w_*(i) \| P_i(x) - x \|^2, \tag{3}$$

is convex, continuously differentiable and it satisfies

$$\lim_{p \to \infty} \nabla\phi_*(x^{k_p}) = 0. \tag{4}$$

Proof. Observe that $\phi_*(x)$ is a convex combination of the convex and continuously differentiable functionals $d_{Q_i}^2(\cdot)$. Therefore, ϕ_* is convex, continuously differentiable and $\nabla\phi_*(x) = 2(x - \mathrm{P}_{w_*}(x))$, for any $x \in \mathscr{H}$. Define the functionals $\phi_k : \mathscr{H} \to \mathbb{R}_+$ by

$$\phi_k(x) := \sum_{i \in I} w_k(i) \| P_i(x) - x \|^2. \tag{5}$$

Similarly to ϕ_*, each functional ϕ_k is convex, continuously differentiable and $\nabla\phi_k(x) = 2(x - \mathrm{P}_{w_k}(x))$. Since, for each $k \in \mathbb{N}$, we have $\| \nabla\phi_k(x^k) \| = 2a_k$ and because $\lim_{k \to \infty} a_k = 0$ by Lemma 4, we obtain

$$\lim_{k \to \infty} \| \nabla\phi_k(x^k) \| = 0. \tag{6}$$

Finally, observe that

$$\begin{aligned} \| \nabla\phi_*(x^k) - \nabla\phi_k(x^k) \| &\leqslant 2 \sum_{i \in I} | w_*(i) - w_k(i) | d_{Q_i}(x^k) \leqslant 2\rho_k \sum_{i \in I} | w_*(i) - w_k(i) | \\ &\leqslant 2\rho_0 \sum_{i \in I} | w_*(i) - w_k(i) |, \end{aligned}$$

because the sequence $\{\rho_k \mid k \in \mathbb{N}\}$ is nonincreasing. Since the sequence $\{w_{k_p} \mid p \in \mathbb{N}\}$ converges to w_* in $\mathbb{R}^I$, it follows that

$$\lim_{p \to \infty} \| \nabla\phi_*(x^{k_p}) - \nabla\phi_{k_p}(x^{k_p}) \| = 0.$$

Combining this and (6), we obtain (4). □

Now we make the first important step towards our final goal.

Lemma 6. (i) $x^* \in Q$.

(ii) $\lim_{p \to \infty} \phi_*(x^{k_p}) = 0$.

(iii) *For each* $i \in I$, *the sequence* $\{d_{Q_i}(x^{k_p}) \mid p \in \mathbb{N}\}$ *converges to zero.*

Proof. For each $y \in \mathscr{H}$, $\phi_*(y) - \phi_*(x^*) \geqslant \langle \nabla\phi_*(x^{k_p}), y - x^{k_p} \rangle + \langle \nabla\phi_*(x^*), x^{k_p} - x^* \rangle$, where the right-hand side converges to zero as $p \to \infty$ (cf. Lemma 5). This shows that x^* is a global minimizer of ϕ_*. Thus, for each $z \in Q$, $0 \leqslant \phi_*(x^*) \leqslant \phi_*(z) = 0$, that is, $\phi_*(x^*) = 0$. Taking into account (3), this implies that, for each $i \in I$, $d_{Q_i}(x^*) = 0$, i.e., $x^* \in Q_i$ because each Q_i is closed. Hence, $x^* \in Q$, which proves (i). According to (i) and Lemma 3, the sequence $\{x^{k_p} - x^* \mid p \in \mathbb{N}\}$ is bounded. Applying Lemma 5, we get

$$0 \leqslant \phi_*(x^{k_p}) = \phi_*(x^{k_p}) - \phi_*(x^*) \leqslant \langle \nabla\phi_*(x^{k_p}), x^{k_p} - x^* \rangle \leqslant \| \nabla\phi_*(x^{k_p}) \| \; \| x^{k_p} - x^* \|,$$

and letting $p \to \infty$ we obtain (ii) by (4). Now, (iii) follows from (ii) since

$$0 \leqslant d_{Q_i}^2(x^{k_p}) = \| P_i(x^{k_p}) - x^{k_p} \|^2 \leqslant \frac{1}{w_*(i)} \phi_*(x^{k_p}). \qquad \square$$

From the results derived above we deduce the following.

Lemma 7. *There exists a subsequence* $\{s_t \mid t \in \mathbb{N}\}$ *of* $\{k_p \mid p \in \mathbb{N}\}$ *such that, for each* $i \in I$, *the sequence* $\{P_i(x^{s_t}) \mid t \in \mathbb{N}\}$ *converges weakly to* x^*.

Proof. Note that, for each $i \in I$ and for all $p \in \mathbb{N}$,

$$\| P_i(x^{k_p}) \| \leqslant \| P_i(x^{k_p}) - x^{k_p} \| + \| x^{k_p} \| = d_{Q_i}(x^{k_p}) + \| x^{k_p} \|.$$

The sequences $\{x^{k_p} \mid p \in \mathbb{N}\}$ and $\{d_{Q_i}(x^{k_p}) \mid p \in \mathbb{N}\}$ are bounded (cf. Lemmas 3 and 6(iii)). Hence, the sequences $\{P_i(x^{k_p}) \mid p \in \mathbb{N}\}$ are also bounded. Therefore, there exists a subsequence $\{s_t \mid t \in \mathbb{N}\}$ of $\{k_p \mid p \in \mathbb{N}\}$ such that each sequence $\{P_i(x^{s_t}) \mid t \in \mathbb{N}\}$ converges weakly to some point $y^i \in \mathscr{H}$. Since Q_i is convex and closed in $\mathscr{H}$, it is also weakly closed. All points $P_i(x^{s_t})$, $t \in \mathbb{N}$, belong to Q_i. Thus, $y^i \in Q_i$, $i \in I$. According to Lemma 6(iii), the sequence $\{P_i(x^{s_t}) - x^{s_t} \mid t \in \mathbb{N}\}$ converges to zero, implying that the sequences $\{P_i(x^{s_t}) \mid t \in \mathbb{N}\}$ and $\{x^{s_t} \mid t \in \mathbb{N}\}$ have the same weak limit, i.e., $y^i = x^*$ for each $i \in I$. Hence, the sequences $\{P_i(x^{s_t}) \mid t \in \mathbb{N}\}$ converge weakly to x^*. $\square$

Let $\theta_{k_p} := \max_{i \in I} d_{Q_i}(x^{k_p})$, for each $p \in \mathbb{N}$. A simple adaptation of the proof of [10, Lemma 5] shows that if $\{\theta_{s_t} \mid t \in \mathbb{N}\}$ converges to zero and at least one of the conditions (A)–(C) of Theorem 1 holds, then the sequence $\{d_Q(x^{s_t}) \mid t \in \mathbb{N}\}$ converges to zero. From Lemma 6(iii) we have that $\lim_{p \to \infty} \theta_{k_p} = 0$. Thus, we deduce that if any of the conditions (A)–(C) of Theorem 1 is satisfied, the sequence $\{d_Q(x^{s_t}) \mid t \in \mathbb{N}\}$ converges to zero. Using that, we make the next step towards proving Theorem 1.

Lemma 8. *Suppose that any of the conditions* (A)–(C) *of Theorem* 1 *is satisfied. Then, the sequence* $\{x^{s_t} \mid t \in \mathbb{N}\}$ *converges strongly to* x^*.

Proof. Denote by B_t the closed ball in $\mathscr{H}$ with center $P_Q(x^{s_t})$ and radius $d_Q(x^{s_t})$, where P_Q represents the orthogonal projection operator onto the closed convex set Q. For each $t \in \mathbb{N}$ the set $S_t := \bigcap_{q=0}^{t} B_q$ is nonempty because, according to Lemma 3, for each $q = 0, 1, \ldots, t$,

$$\| x^{s_{t+1}} - P_Q(x^{s_q}) \| \leqslant \| x^{s_q} - P_Q(x^{s_q}) \| = d_Q(x^{s_q}),$$

i.e., $x^{s_{t+1}} \in S_t$. Since the sequence of closed convex sets $\{S_t \mid t \in \mathbb{N}\}$ satisfies $S_t \supseteq S_{t+1}$ for all $t \in \mathbb{N}$, Hausdorff's completeness theorem (see, e.g., [2, p.61]) implies that the set $S := \bigcap_{t=0}^{\infty} S_t$ is nonempty. Let $\bar{x}$ be any point in S. Then, for each $t \in \mathbb{N}$,

$$\| x^{s_t} - \bar{x} \| \leqslant \| x^{s_t} - P_Q(x^{s_t}) \| + \| P_Q(x^{s_t}) - \bar{x} \| \leqslant 2 d_Q(x^{s_t}),$$

proving that $\lim_{t \to \infty} \| x^{s_t} - \bar{x} \| = 0$ because $\lim_{t \to \infty} d_Q(x^{s_t}) = 0$, as shown above. Hence, the sequence $\{x^{s_t} \mid t \in \mathbb{N}\}$ converges strongly to $\bar{x}$. Since $\{x^{s_t} \mid t \in \mathbb{N}\}$ also converges weakly to x^*, it follows that $x^* = \bar{x}$, whence $\{x^{s_t} \mid t \in \mathbb{N}\}$ converges strongly to x^*. □

The following result closes a cycle in the proof of Theorem 1.

Lemma 9. (i) *If the sequence* $\{x^k \mid k \in \mathbb{N}\}$ *has a subsequence which converges strongly to* x^*, *then* $\{x^k \mid k \in \mathbb{N}\}$ *converges strongly to* x^*.

(ii) *If any of the conditions* (A)–(C) *of Theorem* 1 *is satisfied, then the sequence* $\{x^k \mid k \in \mathbb{N}\}$ *converges strongly to* $x^* \in Q$.

Proof. Let $\{x^{r_n} \mid n \in \mathbb{N}\}$ be a subsequence of $\{x^k \mid k \in \mathbb{N}\}$ which converges strongly to x^*. The fact $x^* \in Q$ follows from Lemma 6(i). Combining this with Lemma 3, the sequence $\{\| x^k - x^* \| \mid k \in \mathbb{N}\}$ and its subsequence $\{\| x^{r_n} - x^* \| \mid n \in \mathbb{N}\}$ are convergent and

$$\lim_{k \to \infty} \| x^k - x^* \| = \lim_{n \to \infty} \| x^{r_n} - x^* \| = 0.$$

This shows that the sequence $\{x^k \mid k \in \mathbb{N}\}$ converges strongly to x^* and (i) is proven. Now, combining (i) with Lemma 8, we obtain (ii). □

The second cycle of the proof of Theorem 1 is achieved next.

Lemma 10. *If one of the sets* Q_i *is boundedly compact, then* $\{x^k \mid k \in \mathbb{N}\}$ *converges strongly to* x^*.

Proof. Suppose that the set Q_{i_0} is boundedly compact. According to Lemma 7, the sequence $\{P_{i_0}(x^{s_t}) \mid t \in \mathbb{N}\}$ converges weakly to x^* and, thus, it is a bounded sequence in Q_{i_0}. Therefore, it includes a strongly convergent subsequence which we again denote $\{P_{i_0}(x^{s_t}) \mid t \in \mathbb{N}\}$. The strong limit of $\{P_{i_0}(x^{s_t}) \mid t \in \mathbb{N}\}$ is exactly x^*, since strong convergence implies weak convergence and the weak limit is necessarily unique. Note that

$$\| x^{s_t} - x^* \| \leqslant \| x^{s_t} - P_{i_0}(x^{s_t}) \| + \| P_{i_0}(x^{s_t}) - x^* \|,$$

and that both terms on the right-hand side of this inequality converge to zero as $t \to \infty$ (cf. Lemma 6(iii)). Hence, the sequence $\{x^{s_t} \mid t \in \mathbb{N}\}$ converges strongly to x^*. By Lemma 9(i), the proof is complete. □

Since in finite-dimensional Hilbert spaces closed sets are boundedly compact, Lemma 10 implies the following result, which completes the proof of Theorem 1.

Lemma 11. *If* $\mathscr{H}$ *is finite-dimensional, then the sequence* $\{x^k \mid k \in \mathbb{N}\}$ *converges* (*strongly*) *to* x^*.

3. Comments

(I) Theorem 1 shows that almost simultaneous BIP methods in a Hilbert space generate sequences which converge strongly to solutions of the given CFP. This generalizes a previous result of Pierra [15, Theorem 1.1] which shows that, if any of the conditions (A)–(C) is satisfied, the simultaneous unrelaxed (i.e., with all $\lambda_k = 1$) BIP methods with uniformly distributed weight functions generate strongly convergent sequences. It follows from [7, Theorem 2.2] that Theorem 1 also generalizes [8, Theorem 7] which shows that simultaneous BIP methods with a constant weight function and fixed relaxation parameters generate strongly convergent sequences whenever one of the sets Q_i is compact. Also, Theorem 1 with condition (D) can be viewed as a nonsequential counterpart of [8, Theorem 3.2(ii)] which guarantees strong convergence of cyclically controlled sequential unrelaxed BIP methods when one of the sets Q_i is boundedly compact. Theorem 1 with condition (E) is a partial restatement of [4, Theorem 4.4] for which we provide an alternative proof.

(II) Pierra [15, Theorem 1.2] has shown that BIP methods with uniformly distributed weight functions and relaxation parameters determined by a specific periodical rule strongly converge to points in Q whenever one of the conditions (A)–(C) of Theorem 1 is satisfied. Ottavy [14, Theorem 5.1 and Corollary 5.1] improved Pierra's results by showing that strong convergence of BIP method generated sequences in Hilbert spaces can be guaranteed under weakened forms of the conditions (A) and (B) when the weight functions and relaxation parameters are chosen according to the following conditions:

$$\exists \theta \in]0, 1[, \forall k \in \mathbb{N}: x^k \notin Q_i \text{ and } w_k(i) > 0 \Rightarrow w_k(i) \geqslant \theta, \tag{7}$$

$$\exists \epsilon > 0, \forall k \in \mathbb{N}: \epsilon \leqslant \lambda_k \leqslant \frac{1}{\| x^k - \mathrm{P}_{w_k}(x^k) \|^2} \phi_k(x^k). \tag{8}$$

Conditions (7) and (8) do not imply that BIP methods satisfying them are simultaneous or almost simultaneous. Also, almost simultaneous BIP method generated sequences whose strong convergence follows from Theorem 1 may not satisfy (7) and/or (8). For instance, if $m = 2$ and any of the conditions (A)–(E) of Theorem 1 is satisfied, then pseudo-periodical BIP method generated sequences with relaxation parameters satisfying (2) and with weight functions $w_k(1) = 1/k$ and $w_k(2) = 1 - 1/k$, for all $k \in \mathbb{N}$ except for those which are divisible by an arbitrarily fixed integer $p > 1$, are almost simultaneous. They strongly converge to points in Q in spite of the fact that condition (7) (and, eventually, condition (8)) is not satisfied. In some circumstances, conditions (7) and (8) considerably restrict the range of choices for λ_k and w_k and, thus, limit the possibility of speeding up convergence of the BIP procedures (see (III) below). For instance, if $m = 2$, Q_1 and Q_2 are halfspaces determined by hyperplanes which intersect each other in an angle α with $\cos \alpha = 0.9$, and if x^k is an exterior equidistant point of the sets Q_1 and Q_2, then the quantity $\phi_k(x^k)/\| x^k - \mathrm{P}_{w_k}(x^k) \|^2$ involved in (8) does not exceed 1.06, no matter how we choose the weight functions w_k. Thus, it imposes on the relaxation parameters a condition which is more restrictive than (2).

(III) Theorem 1 guarantees strong convergence of BIP generated sequences under lax conditions on the weight functions and relaxation parameters. The freedom of choosing w_k and λ_k at each step k can be exploited to improve the behavior of the BIP generated sequences. The *initial speed of convergence* of a BIP method generated sequence $\{x^k \mid k \in \mathbb{N}\}$ at step k can be defined by the number

$$\Delta_k = \inf_{z \in Q} \{ \| x^k - z \| - \| x^{k+1} - z \| \},$$

see [5]. It indicates how much closer to a solution of the CFP we get by moving from x^k to the next iterate x^{k+1}. By increasing the initial speed of a convergent BIP procedure at each step $k=0, 1,\ldots, h$, one may improve its overall initial behavior since the iterate x^h is closer than the initial iterate x^0 to the solution to which the process converges by at least $\tilde{\Delta}_h = \inf_{z\in Q}\{\| x^0 - z \| - \| x^h - z \|\}$ and $\tilde{\Delta}_h \geqslant \Sigma_{k=0}^{h-1}\Delta_k$.

From Lemmas 2(ii) and 3 we deduce that, if $x^k \notin Q$,

$$\Delta_k \geqslant q_k\lambda_k(2-\lambda_k)\,\| \mathrm{P}_{w_k}(x^k) - x^k \|, \tag{9}$$

where $q_k = 0.5 \inf_{z\in Q} \| x^k - z \|^{-1}$. The right-hand side of (9) achieves its largest value when $\lambda_k = 1$ and w_k is a maximum of the convex function $h(w) = \|\mathrm{P}_w(x^k) - x^k\|$. The following example shows that choosing $\lambda_k = 1$ and w_k a maximum of h (as done in "remotest-set" procedures — see [10]) does not always ensure maximal initial speeds.

Example 12. Let $\mathscr{H} = \mathbb{R}^2$, $m=2$, $Q_1 = \{x \mid x_2 = 0.5x_1\}$ and $Q_2 = \{x \mid x_2 = 5x_1\}$. Clearly, $Q = \{0\}$ and Theorem 1 applies. Fix $x^0 = (6, 6)$. Easy computations show that for $\lambda_0 = 1$ the maximal initial speed at step zero, $\Delta_0 \cong 1.8$, is reached with $w_0 \cong (0.32, 0.68)$. By contrast, $h(w)$ achieves its maximum at $w = (0, 1)$ and the corresponding initial speed for the same λ_0 is 1.5. Even better initial speed (i.e., $\Delta_0 \cong 2.5$) is obtained for w_0 as above and $\lambda_0 = 1.5$.

The reason behind the phenomenon observed in Example 12 is that the right-hand side of (9) is a rough lower bound of Δ_k. By guaranteeing strong convergence of BIP generated sequences with relatively lax restrictions for the weight functions and relaxation parameters, Theorem 1 opens the problem of how to determine more accurate lower bounds for the initial speed of convergence.

Acknowledgements

The authors appreciate the comments of two referees which led to improvement of an earlier version of this work.

References

[1] R. Aharoni and Y. Censor, Block-iterative projection methods for parallel computation to solutions of convex feasibility problems, *Linear Algebra Appl.* **120** (1989) 165–175.

[2] G. Bachman and L. Narici, *Functional Analysis* (Academic Press, New York, 1966).

[3] L.M. Bregman, The relaxation method for finding a common point of convex sets and its application to solution of problems in convex programming, *U.S.S.R. Comput. Math. and Math. Phys.* **7** (1967) 200–217.

[4] D. Butnariu and Y. Censor, On the behavior of a block-iterative projection method for solving convex feasibility problems, *Internat. J. Comput. Math.* **43** (1990) 79–94.

[5] D. Butnariu and A. Mehrez, Convergence criteria for generalized gradient methods of solving locally Lipschitz feasibility problems, *J. Comput. Optim. Appl.* **1** (1992) 307–326.

[6] Y. Censor, On variable block algebraic reconstruction techniques, in: G.T. Herman, A.K. Louis and F. Natterer, Eds., *Mathematical Methods in Tomography*, Lecture Notes in Math. **1497** (Springer, Berlin, 1991) 133–140.

[7] P.L. Combettes and J. Trussell, Methods of successive projections for finding a common point of sets in metric spaces, *J. Optim. Theory Appl.* **67** (1990) 487–507.
[8] A.R. De Pierro and A.N. Iusem, A parallel projection method of finding a common point of a family of convex sets, *Pesquisa Oper.* **5** (1985) 1–20.
[9] S.D. Flåm and J. Zowe, Relaxed outer projections, weighted averages and convex feasibility, *BIT* **30** (1990) 289–300.
[10] L.G. Gubin, B.T. Polyak and E.V. Raik, The method of projections for finding common points of convex sets, *U.S.S.R. Comput. Math. and Math. Phys.* **7** (1966) 1–24.
[11] I. Halperin, The product of projection operators, *Acta Sci. Math.* **23** (1962) 94–99.
[12] A.N. Iusem and A.R. De Pierro, Convergence results for an accelerated Cimmino algorithm, *Numer. Math.* **49** (1986) 347–368.
[13] S. Kayalar and H. Wienert, Error bounds for the method of alternating projections, *Math. Control Signals Systems* **1** (1988) 43–59.
[14] N. Ottavy, Strong convergence of projection-like methods in Hilbert spaces, *J. Optim. Theory Appl.* **56** (1988) 433–461.
[15] G. Pierra, Decomposition through formalization in a product space, *Math. Programming* **28** (1984) 96–115.
[16] M.I. Sezan and H. Stark, Applications of convex projections theory to image recovery in tomography and related areas, in: H. Stark, Ed., *Image Recovery: Theory and Applications* (Academic Press, New York, 1987) 415–462.
[17] N.Z. Shor, *Minimization Methods for Nondifferentiable Functions* (Springer, Berlin, 1985).
[18] J. von Neumann, On rings of operators. Reduction theory, *Ann. of Math.* **50** (1949) 401–485.
[19] S.-J. Yeh and H. Stark, Learning in neural nets using projection methods. *Opt. Comput. Process.* **1** (1991) 47–60.